

ENTERED



IFWO

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/671,589

TIME: 08:59:10

Input Set : N:\Cr3\RULE60\10671589.raw

Output Set: N:\CRF4\08052004\J671589.raw

## SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Schlessinger, Joseph

6 Sap, Jan M.

8 (ii) TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE

9 PHOSPHATASE-ALPHA

11 (iii) NUMBER OF SEQUENCES: 14

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: PENNIE & EDMONDS

15 (B) STREET: 1155 AVENUE OF THE AMERICAS

16 (C) CITY: NEW YORK

17 (D) STATE: NEW YORK

18 (E) COUNTRY: U.S.A.

19 (F) ZIP: 10036

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/10/671,589

C--> 29 (B) FILING DATE: 29-Sep-2003

W--> 35 (C) CLASSIFICATION:

37 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: US/09/280,597

34 (B) FILING DATE: 29-MARCH-1999

38 (A) APPLICATION NUMBER: US 08/015,985

39 (B) FILING DATE: 10-FEB-1993

41 (viii) ATTORNEY/AGENT INFORMATION:

42 (A) NAME: Coruzzi, Laura A.

43 (B) REGISTRATION NUMBER: 30,742

44 (C) REFERENCE/DOCKET NUMBER: 7683-020

46 (ix) TELECOMMUNICATION INFORMATION:

47 (A) TELEPHONE: (212) 790-9090

48 (B) TELEFAX: (212) 869-9741/8864

49 (C) TELEX: 66141 PENNIE

53 (2) INFORMATION FOR SEQ ID NO: 1:

55 (i) SEQUENCE CHARACTERISTICS:

56 (A) LENGTH: 802 amino acids

57 (B) TYPE: amino acid

58 (D) TOPOLOGY: linear

60 (ii) MOLECULE TYPE: protein

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ENTERED

ENTERED

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/671,589

TIME: 08:59:10

Input Set : N:\Crif3\RULE60\10671589.raw

Output Set: N:\CRF4\08052004\J671589.raw

```

64 Met Asp Ser Trp Phe Ile Leu Val Leu Leu Gly Ser Gly Leu Ile Cys
65   1           5           10           15
67 Val Ser Ala Asn Asn Ala Thr Thr Val Ala Pro Ser Val Gly Ile Thr
68           20           25           30
70 Arg Leu Ile Asn Ser Ser Thr Ala Glu Pro Val Lys Glu Glu Ala Lys
71           35           40           45
73 Thr Ser Asn Pro Thr Ser Ser Leu Thr Ser Leu Ser Val Ala Pro Thr
74           50           55           60
76 Phe Ser Pro Asn Ile Thr Leu Gly Pro Thr Tyr Leu Thr Thr Val Asn
77 65           70           75           80
79 Ser Ser Asp Ser Asp Asn Gly Thr Thr Arg Thr Ala Ser Thr Asn Ser
80           85           90           95
82 Ile Gly Ile Thr Ile Ser Pro Asn Gly Thr Trp Leu Pro Asp Asn Gln
83           100          105          110
85 Phe Thr Asp Ala Arg Thr Glu Pro Trp Glu Gly Asn Ser Ser Thr Ala
86           115          120          125
88 Ala Thr Thr Pro Glu Thr Phe Pro Pro Ser Gly Asn Ser Asp Ser Lys
89           130          135          140
91 Asp Arg Arg Asp Glu Thr Pro Ile Ile Ala Val Met Val Ala Leu Ser
92 145          150          155          160
94 Ser Leu Leu Val Ile Val Phe Ile Ile Ile Val Leu Tyr Met Leu Arg
95           165          170          175
97 Phe Lys Lys Tyr Lys Gln Ala Gly Ser His Ser Asn Ser Phe Arg Leu
98           180          185          190
100 Ser Asn Gly Arg Thr Glu Asp Val Glu Pro Gln Ser Val Pro Leu Leu
101           195          200          205
103 Ala Arg Ser Pro Ser Thr Asn Arg Lys Tyr Pro Pro Leu Pro Val Asp
104           210          215          220
106 Lys Leu Glu Glu Glu Ile Asn Arg Arg Met Ala Asp Asp Asn Lys Leu
107 225          230          235          240
109 Phe Arg Glu Glu Phe Asn Ala Leu Pro Ala Cys Pro Ile Gln Ala Thr
110           245          250          255
112 Cys Glu Ala Ala Ser Lys Glu Glu Asn Lys Glu Lys Asn Arg Tyr Val
113           260          265          270
115 Asn Ile Leu Pro Tyr Asp His Ser Arg Val His Leu Thr Pro Val Glu
116           275          280          285
118 Gly Val Pro Asp Ser Asp Tyr Ile Asn Ala Ser Phe Ile Asn Gly Tyr
119           290          295          300
121 Gln Glu Lys Asn Lys Phe Ile Ala Ala Gln Gly Pro Lys Glu Glu Thr
122 305          310          315          320
124 Val Asn Asp Phe Trp Arg Met Ile Trp Glu Gln Asn Thr Ala Thr Ile
125           325          330          335
127 Val Met Val Thr Asn Leu Lys Glu Arg Lys Glu Cys Lys Cys Ala Gln
128           340          345          350
130 Tyr Trp Pro Asp Gln Gly Cys Trp Thr Tyr Gly Asn Ile Arg Val Ser
131           355          360          365
133 Val Glu Asp Val Thr Val Leu Val Asp Tyr Thr Val Arg Lys Phe Cys
134           370          375          380
136 Ile Gln Gln Val Gly Asp Met Thr Asn Arg Lys Pro Gln Arg Leu Ile

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/671,589

TIME: 08:59:10

Input Set : N:\Crff3\RULE60\10671589.raw

Output Set: N:\CRF4\08052004\J671589.raw

```

137 385          390          395          400
139 Thr Gln Phe His Phe Thr Ser Trp Pro Asp Phe Gly Val Pro Phe Thr
140          405          410          415
142 Pro Ile Gly Met Leu Lys Phe Leu Lys Lys Val Lys Ala Cys Asn Pro
143          420          425          430
145 Gln Tyr Ala Gly Ala Ile Val Val His Cys Ser Ala Gly Val Gly Arg
146          435          440          445
148 Thr Gly Thr Phe Val Val Ile Asp Ala Met Leu Asp Met Met His Thr
149          450          455          460
151 Glu Arg Lys Val Asp Val Tyr Gly Phe Val Ser Arg Ile Arg Ala Gln
152 465          470          475          480
154 Arg Cys Gln Met Val Gln Thr Asp Met Gln Tyr Val Phe Ile Tyr Gln
155          485          490          495
157 Ala Leu Leu Glu His Tyr Leu Tyr Gly Asp Thr Glu Leu Glu Val Thr
158          500          505          510
160 Ser Leu Glu Thr His Leu Gln Lys Ile Tyr Asn Lys Ile Pro Gly Thr
161          515          520          525
163 Ser Asn Asn Gly Leu Glu Glu Glu Phe Lys Lys Leu Thr Ser Ile Lys
164          530          535          540
166 Ile Gln Asn Asp Lys Met Arg Thr Gly Asn Leu Pro Ala Asn Met Lys
167 545          550          555          560
169 Lys Asn Arg Val Leu Gln Ile Ile Pro Tyr Glu Phe Asn Arg Val Ile
170          565          570          575
172 Ile Pro Val Lys Arg Gly Glu Glu Asn Thr Asp Tyr Val Asn Ala Ser
173          580          585          590
175 Phe Ile Asp Gly Tyr Arg Gln Lys Asp Ser Tyr Ile Ala Ser Gln Gly
176          595          600          605
178 Pro Leu Leu His Thr Ile Glu Asp Phe Trp Arg Met Ile Trp Glu Trp
179          610          615          620
181 Lys Ser Cys Ser Ile Val Met Leu Thr Glu Leu Glu Glu Arg Gly Gln
182 625          630          635          640
184 Glu Lys Cys Ala Gln Tyr Trp Pro Ser Asp Gly Leu Val Ser Tyr Gly
185          645          650          655
187 Asp Ile Thr Val Glu Leu Lys Lys Glu Glu Glu Cys Glu Ser Tyr Thr
188          660          665          670
190 Val Arg Asp Leu Leu Val Thr Asn Thr Arg Glu Asn Lys Ser Arg Gln
191          675          680          685
193 Ile Arg Gln Phe His Phe His Gly Trp Pro Glu Val Gly Ile Pro Ser
194          690          695          700
196 Asp Gly Lys Gly Met Ile Ser Ile Ile Ala Ala Val Gln Lys Gln Gln
197 705          710          715          720
199 Gln Gln Ser Gly Asn His Pro Ile Thr Val His Cys Ser Ala Gly Ala
200          725          730          735
202 Gly Arg Thr Gly Thr Phe Cys Ala Leu Ser Thr Val Leu Glu Arg Val
203          740          745          750
205 Lys Ala Glu Gly Ile Leu Asp Val Phe Gln Thr Val Lys Ser Leu Arg
206          755          760          765
208 Leu Gln Arg Pro His Met Val Gln Thr Leu Glu Gln Tyr Glu Phe Cys
209          770          775          780

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/671,589

TIME: 08:59:10

Input Set : N:\Crif3\RULE60\10671589.raw

Output Set: N:\CRF4\08052004\J671589.raw

```

211 Tyr Lys Val Val Gln Glu Tyr Ile Asp Ala Phe Ser Asp Tyr Ala Asn
212 785                               790                               795                               800
214 Phe Lys
217 (2) INFORMATION FOR SEQ ID NO: 2:
219     (i) SEQUENCE CHARACTERISTICS:
220         (A) LENGTH: 2409 base pairs
221         (B) TYPE: nucleic acid
222         (C) STRANDEDNESS: double
223         (D) TOPOLOGY: unknown
225     (ii) MOLECULE TYPE: cDNA
227     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
229 ATGGATTCCCT GGTTCATTCT TGTTCTGCTC GGCAGTGGTC TGATATGTGT CAGTGCCAAC      60
231 AATGCTACCA CAGTTGCACC TTCTGTAGGA ATTACAAGAT TAATTAACTC ATCAACGGCA      120
233 GAACCAGTTA AAGAAGAGGC CAAAACCTTCA AATCCAACCTT CTTCACTAAC TTCTCTTTCT      180
235 GTGGCACCAA CATTTCAGCCC AAATATAACT CTGGGACCCA CCTATTTAAC CACTGTCAAT      240
237 TCTTCAGACT CTGACAATGG GACCACAAGA ACAGCAAGCA CCAATTCTAT AGGCATTACA      300
239 ATTTACACAA ATGGAACGTG GCTTCCAGAT AACCAGTTCA CGGATGCCAG AACAGAACCC      360
241 TGGGAGGGGA ATTCCAGCAC CGCAGCAACC ACTCCAGAAA CTTTCCCTCC TTCAGGTAAT      420
243 TCTGACTCGA AGGACAGAAG AGATGAGACA CCAATTATTG CGGTGATGGT GGCCCTGTCC      480
245 TCTCTGTAGT TGATCGTGTT TATTATCATA GTTTTGTACA TGTTAAGGTT TAAGAAATAC      540
247 AAGCAAGCTG GGAGCCATTG CAATTCTTTC CGCTTATCCA ACGGCCGCAC TGAGGATGTG      600
249 GAGCCCCAGA GTGTGCCACT TCTGGCCAGA TCCCCAAGCA CCAACAGGAA ATACCCACCC      660
251 CTGCCCCGTG ACAAGCTGGA AGAGGAAATT AACCAGGAGAA TGGCAGACGA CAATAAGCTC      720
253 TTCAGGGAGG AATTCAACGC TCTCCCTGCA TGTCTTATCC AGGCCACCTG TGAGGCTGCT      780
255 TCCAAGGAGG AAAACAAGGA AAAAAATCGA TATGTAAACA TCTTGCCTTA TGACCACTCT      840
257 AGAGTCCACC TGACACCGGT TGAAGGGGTT CCAGATTCTG ATTACATCAA TGCTTCATTC      900
259 ATCAACGGTT ACCAAGAAAA GAACAAATTC ATTGCTGCAC AAGGACCAA AGAAGAAACG      960
261 GTGAATGATT TCTGGCGGAT GATCTGGGAA CAAAACACAG CCACCATCGT CATGGTTACC      1020
263 AACCTGAAGG AGAGAAAGGA GTGCAAGTGC GCCCAGTACT GGCCAGACCA AGGCTGCTGG      1080
265 ACCTATGGGA ATATTGGGGT GTCTGTAGAG GATGTGACTG TCCTGGTGGA CTACACAGTA      1140
267 CGGAAGTTCT GCATCCAGCA GGTGGGCGAC ATGACCAACA GAAAGCCACA GCGCCTCATC      1200
269 ACTCAGTTCC ACTTTACCAG CTGGCCAGAC TTTGGGGTGC CTTTACCCC GATCGGCATG      1260
271 CTCAAGTTCC TCAAGAAGGT GAAGGCCTGT AACCCTCAGT ATGCAGGGGC CATCGTGGTC      1320
273 CACTGCAGTG CAGGTGTAGG GCGTACAGGT ACCTTTGTCT TCATTGATGC CATGCTGGAC      1380
275 ATGATGCATA CAGAACGGAA GGTGGACGTG TATGGCTTTG TGAGCCGGAT CCGGGCACAG      1440
277 CGTGCCAGA TGGTGCAAAC CGATATGCAG TATGTCTTCA TATACCAAGC CCTTCTGGAG      1500
279 CATTATCTCT ATGGAGATAC AGAACTGGAA GTGACCTCTC TAGAAACCCA CCTGCAGAAA      1560
281 ATTTACAACA AAATCCAGG GACCAGCAAC AATGGATTAG AGGAGGAGTT TAAGAAGTTA      1620
283 ACATCAATCA AAATCCAGAA TGACAAGATG CGGACTGGAA ACCTTCCAGC CAACATGAAG      1680
285 AAGAACCGTG TTTTACAGAT CATTCCATAT GAATTCAACA GAGTGATCAT TCCAGTTAAG      1740
287 CGGGGCGAAG AGAATACAGA CTATGTGAAC GCATCCTTTA TTGATGGCTA CCGGCAGAAG      1800
289 GACTCCTATA TCGCCAGCCA GGGCCCTCTT CTCCACACAA TTGAGGACTT CTGGCGAATG      1860
291 ATCTGGGAGT GGAAATCCTG CTCTATCGTG ATGCTAACAG AACTGGAGGA GAGAGGCCAG      1920
293 GAGAAGTGTG CCCAGTACTG GCCATCTGAT GGAAGTGGTGT CCTATGGAGA TATTACAGTG      1980
295 GAAGTGAAGA AGGAGGAGGA ATGTGAGAGC TACACCGTCC GAGACCTCCT GGTCAACCAAC      2040
297 ACCAGGGAGA ATAAGAGCCG GCAGATCCGG CAGTTCCACT TCCATGGCTG GCCTGAAGTG      2100
299 GGCATCCCCA GTGACGGAAA GGGCATGATC AGCATCATCG CCGCCGTGCA GAAGCAGCAG      2160
301 CAGCAGTCAG GGAACCAACC CATCACCGTG CACTGCAGCG CCGGGGCAGG AAGGACGGGG      2220
303 ACCTTCTGTG CCCTGAGCAC CGTCTGGAG CGTGTGAAAG CAGAGGGGAT TTTGATGTC      2280

```

## RAW SEQUENCE LISTING

DATE: 08/05/2004

PATENT APPLICATION: US/10/671,589

TIME: 08:59:10

Input Set : N:\Crif3\RULE60\10671589.raw

Output Set: N:\CRF4\08052004\J671589.raw

```

305 TTCCAGACTG TCAAGAGCCT GCGGCTACAG AGGCCACACA TGGTCCAGAC ACTGGAACAG      2340
307 TATGAGTTCT GCTACAAGGT GGTGCAGGAG TATATTGATG CATTCTCAGA TTATGCCAAC      2400
309 TTCAAGTAA                                2409
312 (2) INFORMATION FOR SEQ ID NO: 3:
314     (i) SEQUENCE CHARACTERISTICS:
315         (A) LENGTH: 793 amino acids
316         (B) TYPE: amino acid
317         (D) TOPOLOGY: linear
319     (ii) MOLECULE TYPE: protein
321     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
323 Met Asp Ser Trp Phe Ile Leu Val Leu Phe Gly Ser Gly Leu Ile His
324   1           5           10           15
326 Val Ser Ala Asn Asn Ala Thr Thr Val Ser Pro Ser Leu Gly Thr Thr
327           20           25           30
329 Arg Leu Ile Lys Thr Ser Thr Thr Glu Leu Ala Lys Glu Glu Asn Lys
330           35           40           45
332 Thr Ser Asn Ser Thr Ser Ser Val Ile Ser Leu Ser Val Ala Pro Thr
333           50           55           60
335 Phe Ser Pro Asn Leu Thr Leu Glu Pro Thr Tyr Val Thr Thr Val Asn
336   65           70           75           80
338 Ser Ser His Ser Asp Asn Gly Thr Arg Arg Ala Ala Ser Thr Glu Ser
339           85           90           95
341 Gly Gly Thr Thr Ile Ser Pro Asn Gly Ser Trp Leu Ile Glu Asn Gln
342           100          105          110
344 Phe Thr Asp Ala Ile Thr Glu Pro Trp Glu Gly Asn Ser Ser Thr Ala
345           115          120          125
347 Ala Thr Thr Pro Glu Thr Phe Pro Pro Ala Asp Glu Thr Pro Ile Ile
348           130          135          140
350 Ala Val Met Val Ala Leu Ser Ser Leu Leu Val Ile Val Phe Ile Ile
351   145          150          155          160
353 Ile Val Leu Tyr Met Leu Arg Phe Lys Lys Tyr Lys Gln Ala Gly Ser
354           165          170          175
356 His Ser Asn Ser Phe Arg Leu Ser Asn Gly Arg Thr Glu Asp Val Glu
357           180          185          190
359 Pro Gln Ser Val Pro Leu Leu Ala Arg Ser Pro Ser Thr Asn Arg Lys
360           195          200          205
362 Tyr Pro Pro Leu Pro Val Asp Lys Leu Glu Glu Glu Ile Asn Arg Arg
363           210          215          220
365 Met Ala Asp Asp Asn Lys Leu Phe Arg Glu Glu Phe Asn Ala Leu Pro
366   225          230          235          240
368 Ala Cys Pro Ile Gln Ala Thr Cys Glu Ala Ala Ser Lys Glu Glu Asn
369           245          250          255
371 Lys Glu Lys Asn Arg Tyr Val Asn Ile Leu Pro Tyr Asp His Ser Arg
372           260          265          270
374 Val His Leu Thr Pro Val Glu Gly Val Pro Asp Ser Asp Tyr Ile Asn
375           275          280          285
377 Ala Ser Phe Ile Asn Gly Tyr Gln Glu Lys Asn Lys Phe Ile Ala Ala
378           290          295          300
380 Gln Gly Pro Lys Glu Glu Thr Val Asn Asp Phe Trp Arg Met Ile Trp

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/671,589

DATE: 08/05/2004  
TIME: 08:59:11

Input Set : N:\Crf3\RULE60\10671589.raw  
Output Set: N:\CRF4\08052004\J671589.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; Xaa Pos.8,12,21,22,24,25,27,28,30,37,39,47,57,72,77,89,94,95,99,104  
Seq#:9; Xaa Pos.109,111,115,116,124,125,131,133,135,137,138,139,143,144,153  
Seq#:9; Xaa Pos.155,170,174,176,179,180,181,182,183,186,205,211,212,214,215  
Seq#:9; Xaa Pos.217,222,227,230,232,240,244,247  
Seq#:14; Xaa Pos.10,20,21,22,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38  
Seq#:14; Xaa Pos.39,40,41,42,43,55,67,70,87,95,96,97,99,100,103,105,106,112  
Seq#:14; Xaa Pos.114,115,116,120,121,123,126,127,128,133,137,138,139,141  
Seq#:14; Xaa Pos.143,149,151,182,186,188,195,196,197,198,199,200,201,202  
Seq#:14; Xaa Pos.205,209,212,213,214,218,222,224,228,229,244,247,254,257  
Seq#:14; Xaa Pos.264,265,266,267

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/671,589

DATE: 08/05/2004

TIME: 08:59:11

Input Set : N:\Crf3\RULE60\10671589.raw

Output Set: N:\CRF4\08052004\J671589.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:35 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1) (vi)  
L:839 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0  
M:341 Repeated in SeqNo=9  
L:1145 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0  
M:341 Repeated in SeqNo=14